

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
30 November 2006 (30.11.2006)

PCT

(10) International Publication Number
WO 2006/125346 A1

(51) International Patent Classification:
G10L 15/26 (2006.01)

(74) Agent: SHANGHAI PATENT & TRADEMARK LAW
OFFICE, LLC; 435 Guiping Road, 200233 Shanghai
(CN).

(21) International Application Number:
PCT/CN2005/000745

(81) Designated States (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(22) International Filing Date: 27 May 2005 (27.05.2005)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (*for all designated States except US*): INTEL CORPORATION [US/US]; 2200 Mission College Boulevard, Santa Clara, CA 95052 (US).

(84) Designated States (*unless otherwise indicated, for every kind of regional protection available*): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(72) Inventors; and

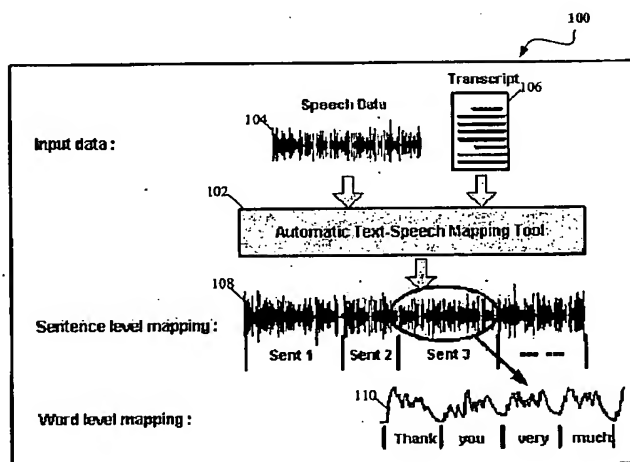
(75) Inventors/Applicants (*for US only*): YEUNG, Minerva [US/US]; 969 Sutter Avenue, Sunnyvale, CA 94086 (US). DU, Robert [CN/CN]; 22nd Floor, ShanghaiMart Tower, No.2299 Yan'An Road(West), 200336 Shanghai (CN). LI, Nan, N. [CN/CN]; 22nd Floor, ShanghaiMart Tower, No.2299 Yan'An Road(West), 200336 Shanghai (CN). WU, Bian [CN/CN]; 22nd Floor, ShanghaiMart Tower, No.2299 Yan'An Road(West), 200336 Shanghai (CN).

Published:

— with international search report

[Continued on next page]

(54) Title: AUTOMATIC TEXT-SPEECH MAPPING TOOL



(57) Abstract: A text-speech mapping method. Silence segments for incoming speech data are obtained. Incoming transcript data is preprocessed. The incoming transcript data comprises a written document of the speech data. Possible candidate sentence endpoints based on the silence segments are found. A best match sentence endpoint is selected based on a forced alignment score. The next sentence is set to begin immediately after the current sentence endpoint, and the process of finding candidate sentence endpoints, selecting the best match sentence endpoint, and setting the next sentence is repeated until all sentences for the incoming speech data are mapped. The process is repeated for each mapped sentence to provide word level mapping.